

# Pemodelan Angka Putus Sekolah Usia SMP Menggunakan Metode Regresi Nonparametrik Spline di Papua

Latifatul Mubarakah, I Nyoman Budiantara, dan Madu Ratna

Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

e-mail: i\_nyoman.b@statistika.its.ac.id

**Abstrak—** Pendidikan yang bermutu dan wajib belajar 12 tahun merupakan prioritas bangsa Indonesia. Pada kenyataannya partisipasi masyarakat dalam pendidikan masih sangat rendah atau angka putus sekolah (APtS) masih relatif tinggi. Berdasarkan Badan Pusat Statistik 2013, APtS di Papua memasuki 10 besar di Indonesia, dengan rata-rata APtS 2,26% yang lebih besar dari rata-rata APtS nasional 2,21%. Salah satu upaya yang dilakukan untuk mengatasi masalah ini adalah menentukan faktor-faktor yang berpengaruh terhadap angka putus sekolah usia SMP di Papua menggunakan regresi nonparametrik spline linier. Pendekatan regresi nonparametrik spline untuk memodelkan APtS di Papua karena pola data pada penelitian ini tidak membentuk suatu pola tertentu. Pada penelitian ini, didapatkan model terbaik yang memiliki nilai GCV minimum dengan tiga knot. Regresi spline linier menghasilkan  $R^2$  sebesar 98,648%, dengan semua variabel signifikan yaitu persentase penduduk miskin ( $x_1$ ), laju ekonomi ( $x_2$ ), rasio guru murid ( $x_3$ ), rasio sekolah murid ( $x_4$ ), dan APtS usia SD ( $x_5$ ).

**Kata Kunci—** APtS, GCV, Regresi Nonparametrik, Spline, Titik Knot.

## I. PENDAHULUAN

Papua merupakan provinsi terluas di Indonesia yang terletak di bagian paling timur Indonesia. Pendidikan merupakan salah satu faktor penting dalam upaya pengembangan Sumber Daya Manusia (SDM). Mengingat hal tersebut, peningkatan mutu pendidikan sebagai upaya pengembangan SDM tentu sangat penting dilakukan. Salah satu parameter keberhasilan dalam peningkatan mutu pendidikan adalah menuntaskan Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM) mutu pendidikan untuk mencapai angka sebesar 95% [1]. Pendidikan di provinsi Papua belum maksimal berdasarkan jenjang pendidikan formal khususnya pada jenjang pendidikan menengah. Hal ini dapat dilihat dari APK dan APM pada buku terbitan Kementerian Pendidikan dan Kebudayaan pada tahun 2013. APK provinsi Papua untuk usia Sekolah Menengah Pertama (SMP) sebesar 82,07% dan APM sebesar 62,91%, dapat dikatakan persentase APK dan APM usia SMP di Papua masih sangat rendah karena masih jauh dengan angka 95%. Besar kecilnya persentase APK dan APM sangat erat

hubungannya dengan Angka Putus Sekolah. Berdasarkan data Badan Pusat Statistik tahun 2010 Papua mempunyai presentase rata-rata Angka Putus Sekolah dan Kemiskinan paling tinggi di Indonesia dengan nilai 33,79%. Sementara itu, untuk Angka Putus Sekolah usia SMP di provinsi Papua masuk ke dalam 10 provinsi dengan angka putus sekolah tertinggi, dengan nilai Angka Putus Sekolah di Papua mencapai 2,26% yang di atas rata-rata nasional sebesar 2,21%. Hal ini perlu adanya penelitian lebih lanjut mengenai faktor-faktor yang mempengaruhi angka putus sekolah di provinsi Papua.

Berbagai penelitian terkait APtS diantaranya tentang analisis faktor-faktor penyebab APtS usia SMA di Jawa Timur [2]. Berdasarkan penelitian yang dilakukan oleh BPS tentang APtS di Indonesia secara umum masalah utamanya adalah masalah ekonomi keluarga. Menurut [3], regresi nonparametrik merupakan regresi yang sangat fleksibel dalam memodelkan pola data. [4] melakukan penelitian tentang faktor-faktor yang mempengaruhi pengangguran terbuka di Aceh dengan menggunakan regresi nonparametrik spline. Berdasarkan penjelasan di atas, masih belum terdapat penelitian yang mengkaji APtS usia SMP di Papua dan faktor-faktor yang mempengaruhinya dengan menggunakan regresi nonparametrik spline. Metode spline dipilih karena kelebihanannya yang dapat digeneralisasikan pada pemodelan statistika yang kompleks dan rumit. Penelitian ini diharapkan mampu menghasilkan model regresi dan faktor-faktor yang mempengaruhi APtS usia SMP di Papua.

## II. TINJAUAN PUSTAKA

### A. Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna [5]. Statistika deskriptif merupakan analisis yang memberikan rangkuman atau ringkasan yang sangat sederhana tentang sampel dan pengamatan yang telah dibuat. Rangkuman tersebut dapat berupa kuantitatif, atau secara visual yaitu agar lebih mudah dalam memahami grafik. Statistika deskriptif dapat dibedakan menjadi dua jenis yaitu ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan data yang umumnya digunakan adalah rata-rata, median, dan modus. Ukuran penyebaran data yang umumnya digunakan adalah keragaman dan *range*.

### B. Regresi Nonparametrik Spline

Regresi nonparametrik merupakan suatu analisis statistika untuk mengetahui hubungan antara variabel dependen dengan variabel independen tanpa harus mengetahui bentuk fungsinya. Regresi nonparametrik merupakan analisis regresi yang sangat fleksibel dalam memodelkan pola data [3]. Secara umum model regresi nonparametrik dapat disajikan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i ; i = 1, 2, 3, \dots, n \quad (1)$$

dengan  $y_i$  merupakan variabel respon (dependen),  $x_i$  merupakan variabel prediktor (independen),  $f(x_i)$  merupakan fungsi regresi nonparametrik serta  $\varepsilon_i$  adalah *error* yang berdistribusi normal, independen dengan *mean* nol dan *varians*  $\sigma^2$ . Regresi spline merupakan potongan polinomial yang mempunyai sifat lebih fleksibilitas dari polinomial biasa sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik lokal dari data. Secara umum fungsi spline  $f(x_i)$  berorde  $q$  dengan titik knot  $K_1, K_2, \dots, K_k$  dapat dituliskan sebagai berikut.

$$f(x_{ji}) = \sum_{l=0}^q \beta_{lj} x_{ji}^l + \sum_{k=1}^r \beta_{(q+k)j} (x_{ji} - K_{kj})_+^q + \varepsilon_i \quad (2)$$

$$\text{dengan } (x_{ji} - K_{kj})_+^q = \begin{cases} (x_{ji} - K_{kj})^q, & x_{ji} \geq K_{kj} \\ 0, & x_{ji} < K_{kj} \end{cases}$$

estimasi parameter pada persamaan (2) dapat diuraikan seperti pada persamaan (3) sebagai berikut.

$$\hat{\beta} = (X'X)^{-1}X'y \quad (3)$$

Regresi nonparametrik spline mempunyai kelebihan dalam mengatasi pola data yang ditunjukkan pada kurva regresi dengan naik atau turun yang tajam dengan bantuan titik-titik knot, serta kurva yang akan dihasilkan akan relatif *smooth* (mulus) [6].

### C. Pemilihan Titik Knot

Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku fungsi pada interval yang berlainan [7]. Salah satu metode yang digunakan untuk pemilihan titik knot adalah *generalized Cross Validation* (GCV). Titik knot yang optimal diperoleh dari GCV yang paling minimum. Metode GCV didefinisikan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_k) = \frac{MSE(K_1, K_2, \dots, K_k)}{(n^{-1} \text{tr}[I - A(K_1, K_2, \dots, K_k)])^2} \quad (4)$$

Dengan  $MSE(K_1, K_2, \dots, K_k) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ,  $K_1, K_2, \dots, K_k$  adalah titik knot dan matriks  $A(K_1, K_2, \dots, K_k)$  diberikan pada persamaan berikut.

$$\hat{y} = A(K_1, K_2, \dots, K_k)\hat{y} \quad (5)$$

### D. Pengujian Parameter Model

Uji parameter dilakukan untuk mengetahui apakah suatu variabel memberikan pengaruh yang signifikan dalam model. Uji untuk parameter dapat diuji secara serentak dan uji secara individu.

#### 1. Uji Serentak

Uji serentak adalah uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor yang dimasukkan ke dalam model memberikan pengaruh secara bersama-sama. Hipotesis yang digunakan untuk pengujian secara serentak adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{q+r} = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0 ; j = 1, 2, \dots, q + r$$

Statistik Uji yang digunakan adalah

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (6)$$

Tabel 1 Analysis of Variance Model Regresi

Sumber Variasi	Degree of Freedom (df)	Sum of Square (SS)	Mean Square (MS)	F Hitung
Regresi	$(q + r)$	$\hat{\beta}'X'y - n\bar{y}^2$	$\frac{SSR}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Error	$n - (q + r) - 1$	$y'y - \hat{\beta}'X'y$	$\frac{SSE}{df_{error}}$	
Total	$n - 1$	$y'y - n\bar{y}^2$		

Tolak  $H_0$  apabila jika  $F_{hitung} > F_{a, q+r, n-(q+r)-1}$  yang berarti bahwa model signifikan [8].

#### 2. Uji Individu

Uji individu adalah uji signifikansi masing-masing variabel prediktor secara individual. Hipotesis yang digunakan dalam uji individu adalah sebagai berikut.

$$H_0: \beta_j = 0 ; j = 1, 2, \dots, q + r$$

$$H_1: \beta_j \neq 0 ; j = 1, 2, \dots, q + r$$

Adapun statistik uji yang digunakan adalah

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (7)$$

dengan:

$$\hat{\beta}_j = \text{penaksir parameter ke-}j$$

$$SE(\hat{\beta}_j) = \text{standart error } \hat{\beta}_j = \sqrt{\text{var}(\hat{\beta}_j)} \text{ dan } \text{var}(\hat{\beta}_j) \text{ sebagai berikut.}$$

$$\text{var}(\hat{\beta}_j) = \text{MSE}(X'X)^{-1}$$

Tolak  $H_0$  apabila nilai  $|t_{hitung}| > t_{\frac{\alpha}{2}, n-(q+r)-1}$  yang berarti bahwa parameter berpengaruh secara signifikan terhadap model [8].

### E. Pengujian Asumsi Residual

Pengujian asumsi residual (*Goodnes of fit*) dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yang identik, independen, dan berdistribusi normal (IIDN).

#### 1. Asumsi Residual Identik

Asumsi identik terpenuhi apabila varians antar residual homogen yakni  $\sigma^2$  dan tidak terjadi heteroskedastisitas [9]. Asumsi residual identik dapat diketahui dengan menggunakan uji Glejser. Adapun hipotesis yang digunakan dalam pengujian Glejser adalah sebagai berikut.

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah sebagai berikut:

$$F_{hitung} = \frac{\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 / (q+r-1)}{\sum_{i=1}^n (|e_i| - |\bar{e}|)^2 / n - (q+r)} \quad (8)$$

Tolak  $H_0$  apabila  $F_{hitung} > F_{(a, (q+r)-1, n-(q+r))}$  atau  $p\text{-value} < \alpha$ . Dimana  $n$  merupakan jumlah observasi dan  $(q + r)$  adalah banyaknya parameter.

#### 2. Asumsi Residual Independen

Asumsi kedua yang harus dipenuhi adalah asumsi residual independen atau tidak terdapat korelasi antar residual yang ditunjukkan oleh nilai kovarian antara  $\varepsilon_i$  dan  $\varepsilon_j$  sama dengan nol. Adapun persamaan ACF sebagai berikut [10].

$$\rho_k = \frac{\sum_{t=k+1}^{n-k} (e_{t-k} - \bar{e})(e_t - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (9)$$

Adapun *confident interval* untuk  $\rho_k$  sebagai berikut:

$$-t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_k) < \rho_k < t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_k) \quad (10)$$

Dengan  $SE(\hat{\rho}_k)$  sebagai berikut.

$$SE(\hat{\rho}_k) = \sqrt{\frac{1}{n}(1 + 2\hat{\rho}_1^2 + \dots + 2\hat{\rho}_m^2)} \quad (11)$$

Apabila terdapat  $\rho_k$  yang keluar dari batas signifikansi, maka dapat dikatakan asumsi independen tidak terpenuhi atau terjadi autokorelasi. Sedangkan apabila tidak terdapat  $\rho_k$  yang keluar dari batas signifikansi, maka asumsi independen terpenuhi.

### 3. Asumsi Residual Normal

Residual dari model regresi harus mengikuti distribusi normal dengan *mean* nol dan varians  $\sigma^2$ . Uji asumsi distribusi normal dapat dilakukan menggunakan uji *Kolmogorov-smirnov* [11]. Adapun uji hipotesis dari pengujian *Kolmogorov-smirnov* adalah sebagai berikut:

$H_0$ : Residual mengikuti distribusi normal

$H_1$ : Residual tidak mengikuti distribusi normal

Adapun statistik uji yang digunakan dalam pengujian *Kolmogorov-smirnov* adalah sebagai berikut:

$$Z_{hitung} = \sup_x |F_n(x) - F_0(x)| \quad (12)$$

Tolak  $H_0$  apabila  $Z_{hitung} > Z_\alpha$  atau  $p - \text{value} < \alpha$ .

### F. Angka Putus Sekolah

Anak putus sekolah merupakan suatu keadaan dimana mencerminkan anak-anak usia sekolah yang sudah tidak bersekolah atau tidak menamatkan suatu jenjang pendidikan tertentu. Angka Putus Sekolah (APT<sub>S</sub>) didefinisikan sebagai perbandingan antara jumlah murid putus sekolah pada jenjang pendidikan tertentu (SD, SMP, dan SMA) dengan jumlah murid pada jenjang pendidikan tertentu dinyatakan dalam presentase. Hasil dari perhitungan APT<sub>S</sub> ini digunakan untuk mengetahui banyaknya siswa putus sekolah di suatu jenjang pendidikan tertentu pada suatu wilayah tertentu. Semakin tinggi APT<sub>S</sub> berarti semakin banyak siswa yang putus sekolah di suatu jenjang pendidikan pada suatu wilayah [12].

## III. METODOLOGI PENELITIAN

### A. Sumber Data dan Variabel Penelitian

Data yang digunakan merupakan data sekunder yang diambil dari Badan Pusat Statistik (BPS) Provinsi Papua dan Susenas Kor tahun 2013. Unit observasi yang digunakan adalah kabupaten/kota yang ada di Papua kecuali Intan Jaya. Variabel-variabel yang digunakan terangkai dalam Tabel 2.

Tabel 2. Variabel Penelitian

Variabel	Keterangan
Y	Angka Putus Sekolah Usia SMP
X <sub>1</sub>	Persentase Penduduk Miskin
X <sub>2</sub>	Laju pertumbuhan ekonomi
X <sub>3</sub>	Rasio Guru Terhadap Murid
X <sub>4</sub>	Rasio Sekolah Terhadap murid Usia SMP
X <sub>5</sub>	APT <sub>S</sub> usia SD

### B. Langkah Analisis

Langkah analisis yang dilakukan adalah sebagai berikut.

1. Menganalisis masing-masing variabel dengan menggunakan analisis statistika deskriptif untuk mengetahui karakteristik dari masing-masing kabupaten/kota di Papua.

2. Membuat *scatter plot* antara jumlah angka putus sekolah usia SMP dan sederajat di Papua dengan masing-masing variabel yang diduga berpengaruh untuk mengetahui bentuk pola data. Apabila terdapat komponen parametrik dan nonparametrik, maka digunakan pendekatan regresi Semiparametrik.
3. Memodelkan data dengan pendekatan Spline satu, dua, tiga knot, dan kombinasi knot.
4. Memilih titik knot optimal berdasarkan nilai GCV yang paling minimum.
5. Mendapatkan titik knot optimal.
6. Menguji signifikansi parameter regresi spline secara serentak.
7. Menguji uji parameter regresi spline secara parsial.
8. Melakukan asumsi residual.
9. Mengintrepetasikan model dan menarik kesimpulan.

## IV. HASIL DAN PEMBAHASAN

### A. Karakteristik APT<sub>S</sub> di Papua

Partisipasi sekolah di Papua tergolong sangat rendah. Pada tahun 2011, rata-rata lama sekolah di Papua mencapai 6,99 tahun yang artinya, rata-rata anak di Papua menamatkan pendidikan hanya sampai di tingkat dasar saja. Sehingga besar kemungkinan bahwa tingkat putus sekolah di Papua cenderung tinggi. Angka putus sekolah di Papua disebabkan oleh beberapa faktor salah satunya faktor ekonomi. Berikut merupakan gambaran keadaan Angka Putus Sekolah dan faktor-faktor yang diduga berpengaruh terhadap APT<sub>S</sub>. Karakteristik APT<sub>S</sub> dan lima variabel prediktor yang diduga berpengaruh disajikan pada Tabel 3.

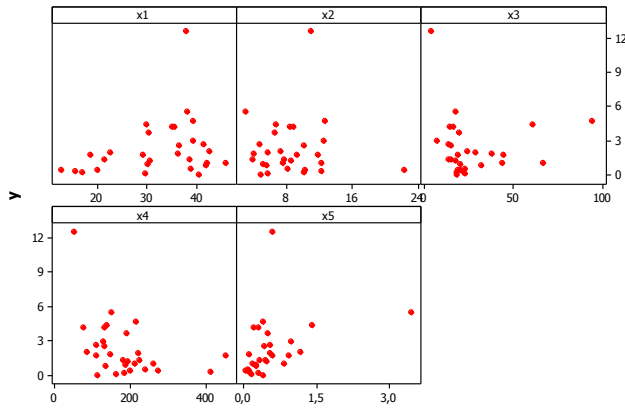
Tabel 3. Deskripsi APT<sub>S</sub> dan Faktor yang Diduga Berpengaruh

Variabel	Rata-rata	Varians	Minimum	Maksimum
y	2,34	6,33	0,06	12,55
x <sub>1</sub>	32,19	86,33	12,95	45,92
x <sub>2</sub>	8,662	15,26	3,05	22,39
x <sub>3</sub>	26,40	384,84	3,72	93,70
x <sub>4</sub>	186,50	7945,30	51,60	455,60
x <sub>5</sub>	0,57	0,43	0,06	3,45

Tabel 3 menjelaskan bahwa rata-rata APT<sub>S</sub> usia SMP di Papua tahun 2013 sebesar 2,34 persen dengan varians sebesar 6,33 persen. APT<sub>S</sub> terendah terdapat pada kabupaten Biak Numfor sebesar 0,06 persen. Sedangkan APT<sub>S</sub> tertinggi terdapat pada kabupaten Nduga sebesar 12,55 persen. Hal ini berarti bahwa APT<sub>S</sub> dari 29 kabupaten di Papua pada tahun 2013 mencapai 0,06 persen hingga 12,55 persen. Secara teori laju pertumbuhan ekonomi berpengaruh terhadap APT<sub>S</sub>. Apabila laju pertumbuhan ekonomi rendah maka APT<sub>S</sub> cenderung tinggi. Pertumbuhan laju ekonomi di Papua memiliki rata-rata sebesar 8,662. Varians dari laju pertumbuhan ekonomi di Papua pada tahun 2013 sebesar 15,26. Pertumbuhan ekonomi tertinggi di Papua pada tahun 2013 dimiliki oleh kabupaten Mimika sebesar 22,39. Kabupaten dengan laju pertumbuhan terendah di Papua pada tahun 2013 dimiliki oleh kabupaten Puncak Jaya dengan laju pertumbuhan ekonomi sebesar 3,05. Begitu pula untuk deskripsi mengenai variabel lain yang dapat dilihat pada Tabel 3.

### B. Pola Hubungan Antara APtS dengan Faktor-Faktor yang Diduga Berpengaruh

Pola hubungan bertujuan untuk melihat pendekatan yang sesuai untuk memodelkan APtS di Papua. Penelitian ini terdiri dari 5 variabel yang diduga berpengaruh terhadap APtS di Papua. Oleh karena itu, didapatkan 5 *scatterplot* dari masing-masing pola hubungan antara variabel respon dan variabel prediktor. Berikut merupakan kelima pola hubungan tersebut.



Gambar 1. Pola Hubungan Antara APtS dengan Faktor yang Diduga Berpengaruh

Gambar 1. Menunjukkan bahwa hubungan antara APtS dengan kelima variabel prediktor tidak membentuk pola tertentu. Sehingga metode yang digunakan untuk memodelkan APtS adalah regresi nonparametrik spline.

### C. Pemilihan Titik Knot Optimum

Pada pemodelan regresi nonparametrik spline dikenal adanya titik knot. Titik knot adalah suatu titik pertemuan dua pola yang berbeda. Pemilihan titik knot yang optimal berdasarkan nilai GCV yang minimum. Perhitungan GCV terdiri dari empat bagian yaitu GCV untuk satu knot, dua knot, tiga knot, dan kombinasi knot. Ringkasan nilai GCV minimum disajikan pada Tabel 4.

Tabel 4. Nilai GCV Terkecil dari Masing-masing Knot

Jumlah Knot	GCV
1 Titik Knot	2,283
2 Titik Knot	2,043
3 Titik Knot	1,319*)
Kombinasi (33313)	1,625

\*) Nilai GCV minimum

Nilai GCV minimum dari model regresi nonparametrik spline menggunakan knot satu, dua, tiga, dan kombinasi titik knot dipilih nilai GCV yang paling minimum yaitu model regresi nonparametrik spline linier dengan tiga titik knot dengan nilai GCV minimum sebesar 1,220. Titik-titik knot optimum pada masing-masing variabel prediktor sebagai berikut.

$$\begin{aligned}x_1 &= 16,987, 19,006, \text{ dan } 21,024 \\x_2 &= 5,418, 6,602, \text{ dan } 7,786 \\x_3 &= 14,738, 20,247, \text{ dan } 25,756 \\x_4 &= 101,047, 125,785, \text{ dan } 150, 523 \\x_5 &= 0,475, 0,683, \text{ dan } 0,890\end{aligned}$$

Model regresi nonparametrik spline linier dengan tiga knot pada variabel yang diduga berpengaruh terhadap APtS usia SMP sebagai berikut.

$$\begin{aligned}\hat{y} &= 29,66 - 0,04x_1 - 1,67(x_1 + 16,987)^{\frac{1}{3}} + \\&4,68(x_1 - 19,006)^{\frac{1}{3}} - 3,065(x_1 - 21,024)^{\frac{1}{3}} - \\&2,38x_2 + 4,027(x_2 - 5,418)^{\frac{1}{3}} - 1,69(x_2 - 6,602)^{\frac{1}{3}} + \\&0,045(x_2 - 7,786)^{\frac{1}{3}} + 0,037x_3 + 0,38(x_3 - 14,73)^{\frac{1}{3}} - \\&0,92(x_3 - 20,24)^{\frac{1}{3}} + 0,58(x_3 - 25,756)^{\frac{1}{3}} - 0,203x_4 + \\&0,415(x_4 - 101,047)^{\frac{1}{3}} - 0,407(x_4 - 125,785)^{\frac{1}{3}} + \\&0,195(x_4 - 150,523)^{\frac{1}{3}} + 0,568x_5 + 27,866(x_5 - 0,475)^{\frac{1}{3}} - \\&58,774(x_5 - 0,683)^{\frac{1}{3}} + 30,938(x_5 - 0,890)^{\frac{1}{3}}\end{aligned}$$

Model regresi nonparametrik spline linier dengan kombinasi knot menghasilkan nilai  $R^2$  sebesar 98,648% yang artinya kelima variabel prediktor mampu menjelaskan sebesar 98,472% dari model.

### D. Pengujian Parameter Model

Uji parameter dalam regresi nonparametrik spline linier dilakukan untuk mengetahui apakah suatu variabel prediktor memberikan pengaruh yang signifikan dalam model atau tidak. Uji untuk parameter dapat diuji secara serentak dan secara individu. Pertama dilakukan pengujian parameter secara serentak lalu dilakukan pengujian parameter secara individu. Uji parameter secara serentak merupakan uji untuk mengetahui apakah semua variabel prediktor yang dimasukkan ke dalam model memberikan pengaruh secara bersama-sama dengan menggunakan uji  $F$ .

Tabel 5. ANOVA Model Regresi Spline Secara Serentak

Sumber Variasi	DF	SS	MS	F hitung
Regresi	20	168,512	8,426	25,544
Error	7	2,309	0,330	
Total	27	170,821		

Berdasarkan hasil tabel ANOVA pada Tabel 4.6 di atas, diperoleh  $p$  - value sebesar 1,05e-05. Karena nilai  $p$  - value  $< \alpha$  dengan  $\alpha = 0,05$  maka dapat diambil keputusan tolak  $H_0$  yang menyimpulkan bahwa minimal ada satu parameter yang berpengaruh signifikan terhadap model. Selanjutnya dilakukan pengujian parameter secara individu yang digunakan untuk mengetahui signifikansi dari masing-masing variabel prediktor secara individu dengan menggunakan uji  $t$ .

Tabel 6. Hasil Pengujian Parameter Model Secara Individu

Variabel	Parameter	Estimator	$p$ - value	Keterangan
Constant	$\beta_0$	29,663	0,0005	signifikan
	$\beta_1$	-0,041	0,855	tidak sig*
	$\beta_2$	-1,671	0,080	tidak sig*
	$\beta_3$	4,677	0,015	signifikan
	$\beta_4$	-3,065	0,007	signifikan
$x_1$	$\beta_5$	-2,381	0,006	signifikan
	$\beta_6$	4,027	0,006	signifikan
	$\beta_7$	-1,695	0,157	tidak sig*
	$\beta_8$	0,046	0,939	tidak sig*
	$\beta_9$	0,037	0,758	tidak sig*
$x_2$	$\beta_{10}$	0,382	0,128	tidak sig*
	$\beta_{11}$	-0,920	0,011	signifikan
	$\beta_{12}$	0,580	0,005	signifikan
	$\beta_{13}$	-0,202	0,0007	signifikan
	$\beta_{14}$	0,415	0,001	signifikan
$x_3$	$\beta_{15}$	-0,407	0,0005	Signifikan
	$\beta_{16}$	0,195	0,0006	signifikan
	$\beta_{17}$	0,568	0,770	tidak sig*
	$\beta_{18}$	27,866	0,018	signifikan
	$\beta_{19}$	-58,774	0,003	signifikan
$x_4$	$\beta_{20}$	30,938	0,001	signifikan

\*)sig = signifikan

Berdasarkan hasil pengujian parameter secara individu pada Tabel 4.7 dapat diketahui bahwa dari 21 parameter yang ada pada model regresi nonparametrik spline linier, terdapat 7 parameter yang tidak signifikan pada taraf signifikansi 0,05. Namun, secara keseluruhan dari 5 variabel prediktor berpengaruh signifikan terhadap APtS usia SMP di Papua.

#### E. Pemeriksaan Asumsi Residual

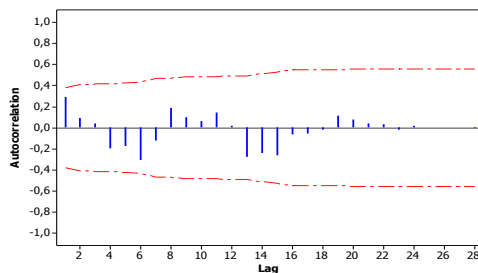
Pemeriksaan asumsi residual dilakukan untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yang identik, independen, dan berdistribusi normal (IIDN). Pemeriksaan asumsi residual identik dilakukan dengan menggunakan uji glejser. Berikut hasil ANOVA pemeriksaan asumsi residual identik dengan uji glejser.

Tabel 7. ANOVA Uji Glejser

Sumber Variasi	DF	SS	MS	F	P-Value
Regresi	20	0,961	0,048	2,076	0,163
Error	7	0,162	0,023		
Total	27	1,123			

Berdasarkan Tabel 7, dapat diketahui bahwa  $p$ -value untuk pemeriksaan asumsi residual identik adalah 0,163. Karena nilai  $p$ -value lebih besar dari taraf signifikansi  $\alpha(0,05)$ , maka dapat disimpulkan bahwa gagal tolak  $H_0$ . Artinya tidak terjadi heteroskedastisitas pada residual atau asumsi residual identik terpenuhi.

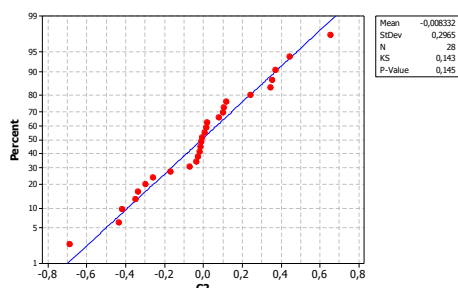
Salah satu cara untuk mengetahui korelasi antar residual bernilai 0 atau tidak adalah dengan cara melihat plot ACF dari residual tersebut. Apabila terdapat minimal satu autokorelasi yang keluar dari batas signifikansi maka terdapat korelasi antar residual. Berikut merupakan plot ACF dari residual model regresi nonparametrik spline.



Gambar 2. Plot ACF Residual

Berdasarkan plot ACF pada Gambar 2, dapat dilihat bahwa tidak ada satupun autokorelasi yang keluar dari batas signifikansi. Sehingga dapat disimpulkan bahwa residual pada model telah memenuhi asumsi independen.

Pemeriksaan asumsi residual berdistribusi normal dilakukan dengan menggunakan uji Kolmogorov-Smirnov. Berikut adalah *probability plot* residual model.



Gambar 3. Plot Distribusi Normal Residual

Secara visual, terlihat bahwa sebaran titik mengikuti garis distribusi normal. Selain itu,  $p$ -value lebih besar dari taraf signifikansi  $\alpha(0,05)$ . Sehingga dapat diambil keputusan bahwa gagal tolak  $H_0$ . Artinya bahwa residual menunjukkan mengikuti asumsi distribusi normal.

#### F. Intrepetasi Model

Model regresi nonparametrik spline terbaik adalah dengan titik knot kombinasi sebagai berikut.

$$\begin{aligned} \hat{y} = & 29,66 - 0,04x_1 - 1,67(x_1 - 16,98)_+^1 + \\ & 4,68(x_1 - 19,006)_+^1 - 3,065(x_1 - 21,024)_+^1 - \\ & 2,38x_2 + 4,027(x_2 - 5,418)_+^1 - 1,69(x_2 - 6,602)_+^1 + \\ & 0,045(x_2 - 7,786)_+^1 + 0,037x_3 + 0,38(x_3 - 14,73)_+^1 - \\ & 0,92(x_3 - 20,24)_+^1 + 0,58(x_3 - 25,756)_+^1 - 0,203x_4 + \\ & 0,415(x_4 - 101,047)_+^1 - 0,407(x_4 - 125,785)_+^1 + \\ & 0,195(x_4 - 150,523)_+^1 + 0,568x_5 + 27,866(x_5 - 0,475)_+^1 - \\ & 58,774(x_5 - 0,683)_+^1 + 30,938(x_5 - 0,890)_+^1 \end{aligned}$$

Model tersebut memiliki lima variabel dimana kelima variabel tersebut berpengaruh signifikan terhadap APtS usia SMP di Papua. Kelima variabel tersebut adalah persentase penduduk miskin ( $x_1$ ), laju pertumbuhan ekonomi ( $x_2$ ), rasio guru terhadap murid, ( $x_3$ ), rasio sekolah terhadap murid usia Smp ( $x_4$ ), dan APtS usia SD ( $x_5$ ). Intrepetasi model bertujuan untuk mengetahui seberapa besar pengaruh masing-masing variabel terhadap APtS usia SMP. Intrepetasi dari model regresi nonparametrik spline terbaik adalah apabila variabel  $x_2$ ,  $x_3$ ,  $x_4$ , dan  $x_5$  diasumsikan konstan, maka pengaruh persentase penduduk miskin terhadap APtS usia SMP sebagai berikut.

$$\begin{aligned} \hat{y} = & -0,04x_1 - 1,67(x_1 - 16,98)_+^1 + 4,68(x_1 - 19,006)_+^1 - \\ & 3,065(x_1 - 21,024)_+^1 \\ = & \begin{cases} -0,04x_1, & x_1 < 16,98 \\ -1,71x_1 + 27,48, & 16,98 \leq x_1 < 19,006 \\ 2,97x_1 - 61,47, & 19,006 \leq x_1 < 21,024 \\ -0,095x_1 + 2,97, & x_1 \geq 21,024 \end{cases} \end{aligned}$$

Pada kondisi persentase penduduk miskin kurang dari 16,98, jika terjadi kenaikan satu satuan maka APtS usia SMP cenderung turun sebesar 0,04 persen. Wilayah yang termasuk dalam interval ini adalah kabupaten Merauke dan kota Jayapura. Sedangkan pada kondisi persentase penduduk miskin diantara 16,98 sampai 19,006, jika terjadi kenaikan satu satuan maka APtS usia SMP cenderung turun sebesar 1,71 persen. Wilayah yang termasuk dalam interval ini adalah kabupaten Jayapura, dan Sarmi.

Apabila pada kondisi persentase penduduk miskin diantara 19,006 sampai 21,004 persen, jika terjadi kenaikan satu satuan maka APtS usia SMP akan cenderung naik sebesar 2,97 persen. Wilayah yang termasuk dalam interval ini adalah kabupaten Mimika.

Sedangkan pada kondisi persentase penduduk miskin berada di atas 21,024 persen, jika terjadi kenaikan satu satuan maka APtS usia SMP cenderung turun sebesar 0,09 persen. Wilayah yang termasuk ke dalam interval ini adalah kabupaten Jayawijaya, Nabire, Yapen Waropen, Biak Numfor, Paniai, Puncak Jaya, Boven Digoel, Mappi, Asmat, Yahukimo, Pegunungan Bintang, Tolikara, Waropen, Supiori, Mamberamo Raya, Nduga, Lanny Jaya, Mamberamo Tengah, Yalimo, Puncak, Dogiyai, dan Deiyai.

Variabel lainnya dapat diinterpretasikan dengan cara yang sama dengan  $x_1$ .

## V. KESIMPULAN DAN SARAN

Faktor-faktor yang mempengaruhi APtS usia SMP di Papua pada tahun 2013 adalah persentase penduduk miskin ( $x_1$ ), laju pertumbuhan ekonomi ( $x_2$ ), rasio guru terhadap murid, ( $x_3$ ), rasio sekolah terhadap murid usia SMP ( $x_4$ ), dan APtS usia SD ( $x_5$ ). Berdasarkan model regresi nonparametrik spline linier yang optimal dengan kombinasi knot didapatkan nilai  $R^2$  sebesar 98,648% dan nilai MSE sebesar 0,329 serta asumsi residual yang telah terpenuhi sehingga dapat dikatakan baik dalam pemodelan.

Saran yang dapat direkomendasikan pada penelitian selanjutnya dan untuk kebijakan pemerintah adalah perlunya menambahkan jumlah knot menjadi empat knot dan diharapkan adanya upaya peningkatan dari pemerintah dalam meningkatkan kualitas penduduk untuk mengurangi APtS usia SMP di provinsi Papua, seperti peningkatan kualitas pendidikan dasar dan mulai memberikan kebijakan untuk meningkatkan laju perekonomian daerah setiap kabupaten di provinsi Papua.

## DAFTAR PUSTAKA

- [1] Rasiyo. (2008). *Pemerataan Pendidikan Belum Tercapai*. Diunduh dari alamat <http://els.bappenas.go.id/upload/klipping/pemerataan%20%20-blm.pdf> pada Minggu, 12 April 2015, 10.00 pm.
- [2] Pradipta, M. (2013). *Pemodelan Angka Putus Sekolah Usia SMA di Jawa Timur dengan Pendekatan Regresi Spline Multivariabel*: Tugas Akhir ITS.
- [3] Eubank, R. (1998). *Spline Smoothing and Nonparametric Regression*. Marcel Dekker, New York.
- [4] Sani, L.N. (2015). *Pemodelan Faktor-faktor yang Mempengaruhi pengangguran Terbuka di Provinsi Aceh dengan Menggunakan Pendekatan Regresi Nonparametrik Spline*: Tugas Akhir ITS.
- [5] Walpole, R. E. (1995). *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama.
- [6] Hardle, Wo., Marlene, M., Stefan, S., dan Axel, W. (2004). *Nonparametric and Semiparametric Models*. Berlin: Springer.
- [7] Budiantara, I. N. (2006). "Regresi Nonparametrik Dalam Statistika". Makalah pembicara Utama pada Seminar Nasional Matematika, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Makasar (UNM), Makasar.
- [8] Draper, N.R. dan Smith, H. (1996). *Applied Regression Analysis, Second Edition*. John Wiley & Sons.
- [9] Gujarati, N. D. (2004). *Basic Econometrics, Fourth Edition*. New York: The McGraw-Hill Companies.
- [10] Wei, W. W. (2006). *Time Series Analysis Univariate and Multivariate Methods 2nd Edition*. New Jersey: Pearson Education.
- [11] Daniel, W.W. (1989). *Applied Nonparametric Statistics*. New York: John Wiley.
- [12] Wakhinuddin. (2009). *Angka Partisipasi dalam Pendidikan*. Diunduh dari alamat <https://wakhinuddin.word-press.com/2009/08/07/angka-partisipasi-dalam-pendidikan/> pada Minggu, 12 April 2015, 09.30 pm.